

Landscape of kinetically trapped binary assemblies

Ranjan V. Mannige*

Molecular Foundry, Lawrence Berkeley National Laboratory, 1 Cyclotron Road, Berkeley, CA, U.S.A.

(Dated: November 16, 2015)

For two-component assemblies, an inherent structure diagram (ISD) is the relationship between set inter-subunit energies and the types of kinetic traps (inherent structures) one may obtain from those energies. It has recently been shown that two-component ISDs are apportioned into regions or plateaux within which inherent structures display uniform features (e.g., stoichiometries and morphologies). Interestingly, structures from one of the plateaux were also found to be robust outcomes of one type of non-equilibrium growth, which indicates the usefulness of the two-component ISD in predicting outcomes of some types of far-from-equilibrium growth. However, little is known as to how the ISD is apportioned into distinct plateaux. Also, while each plateau displays classes of structures that are morphologically distinct, little is known about the source of these distinct morphologies. This article outlines an analytic treatment of the two-component ISD, and shows that the manner in which any ISD is apportioned arises from a single unitless order parameter. Additionally, the analytical framework allows for the characterization of local properties of the trapped structures within each ISD plateau. This work may prove to be useful in the design of novel classes of robust nonequilibrium assemblies.

I. INTRODUCTION

An important goal of molecular design is to facilitate the *robust* self-assembly of elaborate structures from smaller components. Here, “robust” refers to the capacity of the same components starting off in varying environmental conditions to result in the same or similar outcome^{1–6}. The outcome of a robust molecular assembly process is traditionally thought to be the equilibrium (lowest-free-energy) structure, with non-equilibrium structures posing hurdles (kinetic traps) towards that process^{1–11}. Yet, some regimes of far-from-equilibrium growth have also been shown to result in robust outcomes^{12–15}, which indicates novel (non-equilibrium) routes towards building robust molecular assemblies.

While non-equilibrium outcomes in relatively detailed and complex processes have been studied – e.g., in the process of virus capsid assembly⁶ and DNA hybridization¹⁰ – computational limitations prevent more systematic studies of such systems. Alternatively, simpler ordered frameworks formed from two types of components have been useful in systematically distinguishing the outcomes of equilibrium and non-equilibrium regimes for nucleation and growth^{12–19}. This report considers growth post nucleation, where, depending on the rate of growth, both equilibrium and non-equilibrium results can be resolved^{14,15}. For convenience, this report refers to the two component types as “red” and “blue” (as is the convention in Refs. 14 and 15).

Near-equilibrium growth is attained when the rate of growth is so slow that binding is nearly reversible. The first utility of two-component growth is that the outcomes of equilibrium growth are relatively easy to predict. For example, if the heterogeneous energy of interaction is dominantly the lowest in energy, analogous to the components sodium (Na^+) and chloride (Cl^-) in a crys-

tal, the equilibrium outcome is “binary” (otherwise the outcomes dominantly display one or the other subunit). Here, binary indicates that each neighbor is surrounded by neighbors of opposite color. In a two-dimensional (2d) square lattice, such a binary structure resembles a checkerboard structure (Fig. 1a). A convenient metric to characterize binary assemblies is the fraction of blues in the assembly (f_b), which, in the case of a binary solid is 0.5 (like in Na^+Cl^-).

When growing an assembly at far-from-equilibrium rates (i.e., at much greater growth rates), the expected equilibrium outcomes (such as structures displaying $f_b = 0.5$ for the system above) are not the outcomes that one encounters^{12–17,19}. This is because, at particular growth rates and subunit-subunit energetics, kinetic traps involving multiple components within the bulk of the assembly are not permitted to rearrange into equilibrium features before those defects get buried²⁰. Mostly, non-equilibrium outcomes are highly sensitive to solution conditions such as the combined and relative concentration of the two components in solution. However, in some cases, the non-equilibrium outcomes are actually robust to some growth conditions^{12–15}. Presently, while such robust yet non-equilibrium assemblies may present us with new ways to produce novel and controllable materials, an overarching framework for making predictions for such processes is lacking, especially because most existing frameworks for predicting molecular outcomes are based on equilibrium statistical mechanics (see discussions in Refs. 18 and 19).

Refs. 14 and 15, which will be focused on in more detail here, describe a non-equilibrium growth scenario that yields assemblies that are robust to *multiple* solution properties: growth rate and solution stoichiometries. The model assumes the following hierarchy of subunit-subunit interaction energies^{14,15}: the red-blue interaction energy (ϵ_{br}) is the most favorable, the red-red interaction energy (ϵ_{rr}) is negligible and intermediate in strength,

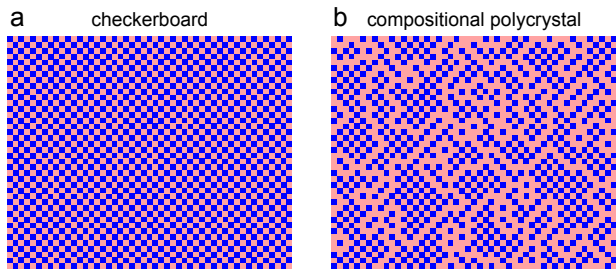


Figure 1. Equilibrium versus nonequilibrium structures obtained from growth. Given two components (“red” and “blue”) whose heterogeneous interaction energy is the lowest (i.e., $\epsilon_{br} < \epsilon_{bb}, \epsilon_{rr}$), near-equilibrium growth results in a “checkerboard” patterned structure with equal ratio of blues and reds ($f_b = 0.5$), shown in (a) for a 2d square lattice (henceforth, the ‘red’ components are shown faded for increased contrast compared to blue components). Interestingly, at a particular interaction energy regime ($\epsilon_{br} < \epsilon_{rr} \equiv 0 < 2|\epsilon_{br}| < \epsilon_{bb}$), while the equilibrium structure remains the checkerboard (a), a new class of non-equilibrium structures – compositional polycrystalline assemblies (b) with a particular f_b (~ 0.364) – form from a range of non-equilibrium growth regimes and red:blue ratios in solution^{14,15} (also shown in Fig. 2a,b). This particular class of polycrystalline structure is but one of a range of possible structures available within the inherent structure diagram (Fig. 2d).

and the blue-blue interaction energy (ϵ_{bb}) is high enough to never permit its occurrence during growth/assembly (all interaction energies in this manuscript are taken to be the equivalent of binding free energies in actual molecules). These stipulations equate to the hierarchy $\epsilon_{br} < \epsilon_{rr} \equiv 0 < 2|\epsilon_{br}| < \epsilon_{bb}$. Subunit growth was simulated on a lattice (Fig. 2a; also see Section IIB), which yielded in Refs. 14 and 15 the following results.

At slow, near-reversible, growth rates (μ is taken as a proxy), the expected equilibrium outcome is obtained ($\mu = -15$ always yields $f_b = 0.5$; Fig. 2b). However, some faster growth rates (about $-2 < \mu < 8$; ‘★’ in Fig. 2b) yield f_b values that deviate from the equilibrium f_b . Yet, these structures are still robust to *both* a range of growth rates and solution stoichiometries^{14,15}. Robustness to growth rate ($\propto \mu$) is evident in the presence of a plateau (★) per curve in Fig. 2b; Robustness to solution stoichiometry is evident in the adherence of all ‘★’ plateaux to single magic number ($f_b \approx 0.364$), regardless of the distinct ratios of red:blue set in in solution ($f_b^s = 0.2, 0.5, 0.8$). The plateau ‘★’ describes a range of growth rates that are neither reversible nor irreversible to the chosen interactions energies: they describe a partially reversible regime of growth where some types of binding are allowed to reverse (e.g., r-r) while some bind irreversibly (e.g., r-b)^{14,15}. The nonequilibrium assemblies may be termed as “compositional polycrystals”^{14,15} whose “grains” are patches of equilibrium regions (checkerboard) and whose boundaries constitute only red components (Fig. 1b). The word “compositional” in this term is meant to indicate that, ignor-

ing color, the framework would be fully occupied and ordered^{14,15}.

Interestingly, these polycrystalline structures obtained via partially reversible growth are also obtained *via* a simpler sampling protocol (Fig. 2c; see Section IIC)^{14,15} that samples for inherent structures^{21–23}. Inherent structures were obtained by a two-state sampling technique that allows only energetically favorable or neutral color switches (Fig. 2c). Inherent structure sampling was applied to a range of random initial configurations at particular points in $(\epsilon_{bb}, \epsilon_{br})$ -space ($\epsilon_{rr} \equiv 0$ with no loss of generality¹⁴). The resulting average f_b as a function of $(\epsilon_{bb}, \epsilon_{br})$, which may be referred to as an inherent structure diagram (ISD), is shown in Fig. 2d. Interestingly, the ISD is apportioned into discrete regions (plateaux) within which the structures all describe nearly identical f_b . The region within the ISD that describes plateau 1 describes structures that are identical to those found from partially reversible growth (★ in Fig. 2b): structures within plateau 1 are compositional polycrystals (Fig. 1b) whose f_b s are identical to those obtained from partially reversible growth. Additionally, all points within plateau 1 exclusively satisfy the interaction energy relationship that was used to grow the non-equilibrium structures (Fig. 2b), namely, $\epsilon_{br} < \epsilon_{rr} \equiv 0 < 2|\epsilon_{br}| < \epsilon_{bb}$. This indicates a substantial connection between the structures available within an inherent structure diagram and the outcomes of some types of non-equilibrium growth; the other plateaux in the inherent structure diagram – each describing potential materials with specific local structural properties (Fig. 5) – may, too, lead to the nonequilibrium growth of other materials with other magic numbers.

While the potential exists for inherent structures to explain the outcomes of two-component out-of-equilibrium assemblies, little is known about what determines the plateaux evident in the ISD. This manuscript describes a pen-and-paper model that will be used to show that each ISD plateau, as well as the properties (local environments) of the inherent structures within each plateau, are consequences of the local connectivity (degree) of the assembly framework (lattice). Specifically, the formalism developed below predicts *i*) the boundaries of potential plateaux in any two-component ISD, and *ii*) local environments of each plateau-specific inherent structure. This report also shows that the local rules developed here can not predict the actual magic ratios (f_b) each plateau displays, as those are a matter of global (not local) lattice connectivity¹⁴.

II. NUMERICAL METHODS DESCRIBED PREVIOUSLY^{14,15}

This section is meant to accompany the discussions in the introduction pertaining to the correspondence between outcomes of non-equilibrium growth and inherent structure sampling.

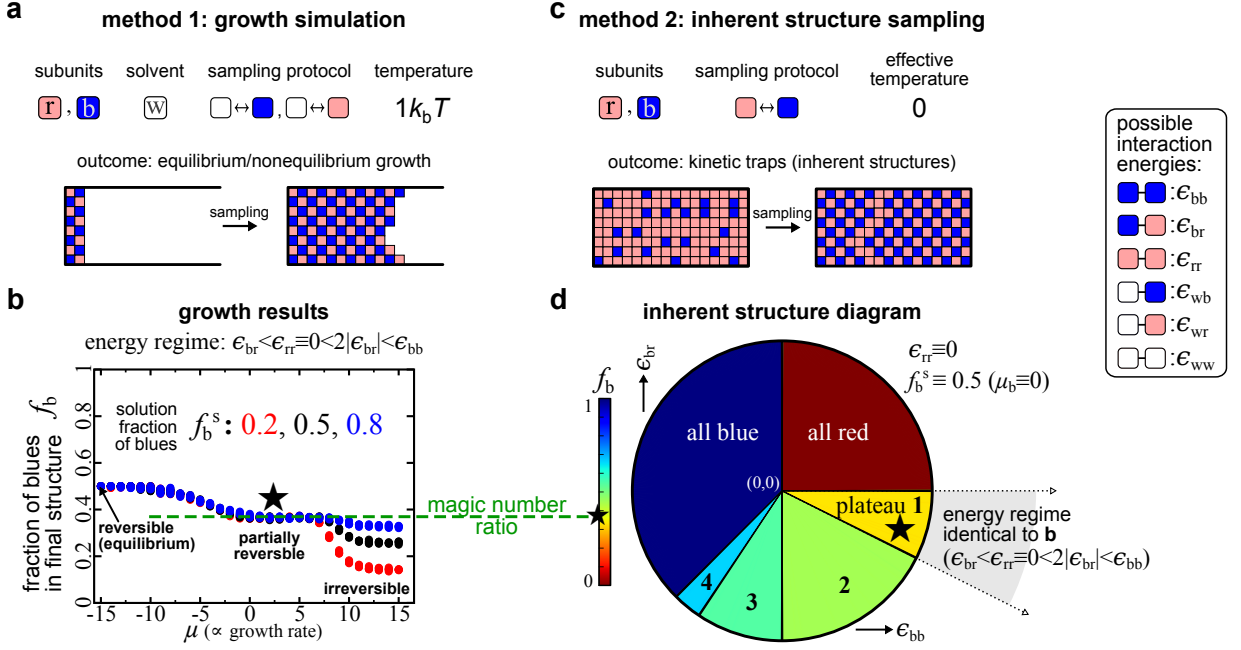
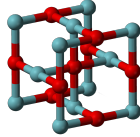


Figure 2. **Previously discussed growth (a,b)¹⁵ and inherent structure sampling (c,d)¹⁴.** As replicated from a previous study¹⁵, Monte Carlo simulations of growth (a; see Section II B) at a certain energy regime result in assemblies whose composition (fractions of blues; f_b) remain robust to both a range of growth rates and relative solution concentrations (\star in b)^{14,15}. Inherent structure sampling of a random two-component lattice model (c; see Section II C) results in uniform regions (plateaux) of interaction space within which inherent structures describe uniform f_b (d)¹⁴. Interestingly, the robust structures obtained from growth at partially reversible rates (\star) are identical to those obtained from plateau 1 of the inherent structure diagram, regardless of a range of starting configurations^{14,15}. The structures represent a class of compositional polycrystals (Fig. 1b).

A. Interaction graphs used in the Monte Carlo sampling

Any arbitrary interaction network (graph) may be used in our sampling procedure below. Particular graphs used here are the periodically connected square lattices in dimension d ($d \in \{1, 2, \dots, 5\}$) and the **nbo** lattice. The **nbo** lattice²⁴ represents the interaction network observed in a metal-organic framework (MOF-2000) described before¹⁴. The periodic cell of the **nbo** graph is as follows:



(**nbo** topology; image source: Wikipedia²⁵)

Note that the higher dimension square lattices ($d > 3$), although unnatural, are useful to understand materials characterized by higher degrees or coordination number.

B. Method 1: Monte Carlo sampling algorithm for nucleated growth (Fig. 2a)¹⁵.

The sluggish dynamics of particles within a solid is a primary cause of kinetic trapping within a growing

multicomponent assembly. Due to this simple physical origin, kinetic trapping from growth can be reproduced by simple physical models that account for this slow dynamic^{12,13,16,17}. This section discusses a lattice model of growth similar to the models used in Refs. 11, 14, 19 and identical to that used in Ref. 15. As sketched in Fig. 2a, each site on the lattice may be either occupied by a subunit/component colored red or blue or by solvent that is colored white. The energy of the system is

$$E = \sum_{i,j}^{\text{edges}} \epsilon_{C(i)C(j)} + \sum_i^{\text{nodes}} \mu_{C(i)}. \quad (1)$$

The first sum iterates through all nearest-neighbor interactions. Here $C(i) = \text{b}, \text{r}$ or w if node i is ‘blue’, ‘red’ or ‘white’, respectively. $\epsilon_{C(i)C(j)}$ is the pairwise free energy of interaction between two adjacent nodes i and j (simply called interaction ‘interaction energy’ henceforth); therefore, there are six types of interaction energies: ϵ_{bb} , ϵ_{br} , ϵ_{rr} , ϵ_{wb} , ϵ_{wr} , and ϵ_{ww} .

The second iterates through all nodes, which represent subunits (b,r) and solvent (w). The chemical potential term $\mu_{C(i)}$ is μ , $-\ln(f_b^s)$ and $-\ln(1 - f_b^s)$ for w, b and r, respectively.

In the absence of pairwise energetic interactions (i.e. in implicit solution), the likelihood that a given site will be white, blue or red is respectively $\{p_w, p_b, p_r\} =$

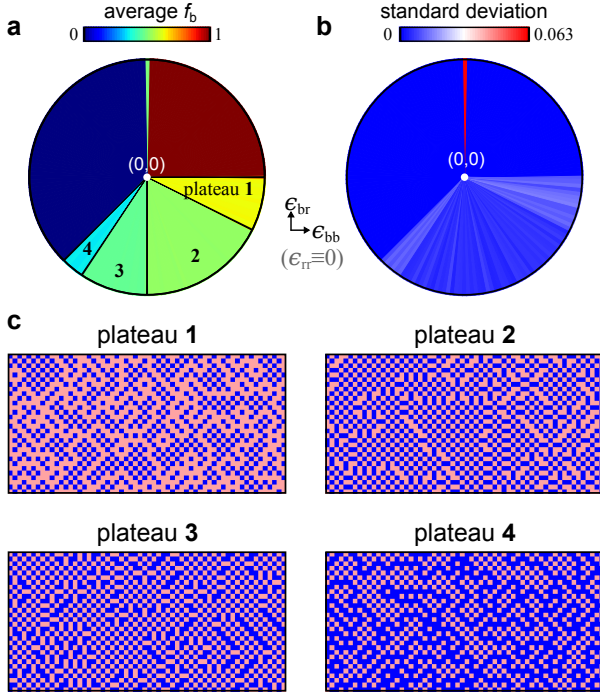


Figure 3. **Each plateau in (a) displays homogeneous f_b (b) and distinct classes of polycrystalline inherent structure (c).** It has already been shown that inherent structures observed in plateau 1 of the inherent structure diagram or ISD (a; Fig. 2d) represent compositional polycrystals with binary domains/grains and all-red grain boundaries¹⁴. Interestingly, all other non-trivial regions of the ISD (plateaux 2 through 4) also describe distinct classes of compositional polycrystals, whose structures are distinguished by distinct grain boundaries (e.g., plateau 4 describes boundaries that are all blue). As indicated by (b), while each plateau describes a range of compositional polycrystals, each structure within the same plateau describes the same f_b .

$$\{e^{-\mu}, f_b^s, 1 - f_b^s\} (1 + e^{-\mu})^{-1}.$$

The Monte Carlo growth simulations were performed in the following manner. The simulation box was 400 sites wide by 40 sites high. The first six columns were populated with the equilibrium checkerboard structure. Our sampling protocol involved selecting a node at random and attempting to change the color of that node. Given a white node, an attempt is made to change its color to either red or blue, and given a red/blue colored site, an attempt is made to change its color to white.

When attempting to switch from white to colored, blue is chosen with probability f_b^s , and red is chosen otherwise. Red-to-blue switches are not allowed, which mimics the idea that binding and unbinding events are the dominant way in which the configurational degrees of freedom evolve.

To maintain detailed balance with respect to the Eqn. 1, the acceptance rates for these moves were as follows (ΔE is the energy change resulting from the pro-

posed move):

$$\begin{aligned} r \rightarrow w &: \min(1, (1 - f_b^s) \exp[-\Delta E]); \\ w \rightarrow r &: \min(1, (1 - f_b^s)^{-1} \exp[-\Delta E]); \\ b \rightarrow w &: \min(1, f_b^s \exp[-\Delta E]); \\ w \rightarrow b &: \min(1, (f_b^s)^{-1} \exp[-\Delta E]). \end{aligned}$$

The implicit solution abundances of red and blue are controlled by the chemical potential term that appears in Eqn. 1 and therefore in the term ΔE . Our choice to insert blue particles with likelihood f_b^s does not by itself result in a thermodynamic bias for one color over the other (because this bias in proposal rate is countered by the non-exponential factors in the acceptance rates). Instead, insertions are biased so that the dynamics of association are consistent with the thermodynamics of the model. For instance, if blue particles are more numerous in solution than red ones, it is physically appropriate to insert blue particles into the simulation box more frequently than red particles. Consider the limit of large positive μ : the ‘solid solution’ that results as the box fills irreversibly with colored particles will have a red:blue stoichiometry equal to that of the notional solution only if blue particles are inserted with likelihood f_b^s . As a technical note, the chemical potential term present in ΔE ends up simply canceling the non-exponential factors in the acceptance rates, but Ref. 15 chose to write acceptance rates as shown in order to make clear which pieces are imposed by thermodynamics, and which pieces are chosen for dynamical reasons. Note that temperature is not defined explicitly, but can be considered to be subsumed into the energetic parameters of the model.

The parameter values used to obtain Fig. 2d in $k_b T$ are $\epsilon_{bb} = 70, \epsilon_{br} = -7, \epsilon_{rr} = 0$, which is replicated from Ref. 15.

C. Method 2: Sampling algorithm to obtain inherent structures (Fig. 2c)¹⁴.

As in Method 1, our energy function follows that of Eqn. 1, with the exception that white (solvent) sites are not allowed (Fig. 2c). Therefore, there are three types of interaction (free) energies: ϵ_{bb} , ϵ_{br} , ϵ_{rr} . The chemical potential term $\mu_{C(i)}$ is 0 and μ_b for red and blue nodes, respectively. These chemical potentials indicate how easy it is to switch to red and blue states based on the relative concentration of reds vs blues in a notional or implicit ‘solution’ (the grand canonical reservoir of colored blocks). Therefore, the fraction of blue nodes in implicit solution is

$$f_b^s = \frac{e^{\mu_b}}{(1 + e^{\mu_b})}. \quad (2)$$

Inherent structure sampling happens in the following fashion. All sites within the lattice are randomly populated with subunits (either red or blue). At each sampling

step, state changes are attempted by randomly selecting a site and then attempting a switch in the subunit's color. If the switch is energetically favorable or neutral, then that state change is kept. This process continues until no state change occurs. While these rules in effect represent sampling at $T = 0$, they are also equivalent to a sampling protocol where temperature is finite (say $T = 1$) and where the interaction energies are multiplied by ∞ .





All of the inherent structure diagrams or ISDs discussed in this paper (Fig. 2c, Fig. 5a-b, Fig. 5a-c) are calculated with $\mu_b \equiv 0$, i.e., relative concentrations of red versus blue subunits in the notional solution are equal ($f_b^s \equiv 0.5$). Changing the relative concentration (by tuning μ_b) effectively shifts or translates the features (plateaux) within the ISD with respect to ϵ_{bb} and ϵ_{rr} without changing the positions of the plateaux relative to each other (see Fig. 3B in Ref. 14).






As calculated previously^{14,15}, each point on the inherent structure diagram obtained in Fig. 2b represents the average fraction of blues (f_b) of a number of inherent structures, calculated using the method shown above, at that particular ($\epsilon_{bb}, \epsilon_{br}$) point.







A theoretical framework is explored below that attempts to chart such inherent structure diagrams (ISDs).







III. THEORY: CHARTING THE INHERENT STRUCTURE DIAGRAM

Possible ISD plateau boundaries from simplex transitions. First, this report shows that borders between plateaux may be surmised by understanding the behavior of color switches within simpler sub-structures of the lattice. Such units of structure are described as “simplices”, which comprise a central node and all of its nearest neighbors. While the term “simplex” is traditionally considered to be extensions of a triangle in a given dimension, this report utilizes the algebraic topology definition of a simplex (abstract simplicial complex), in which the word “simplex” means a finite (but still useful and basic) set of vertices. For the purpose of this paper, a simplex is a convenient word to be used repeatedly instead of the longer yet appropriate term “central node with all its neighboring nodes”.

While simplices can be described for any regular graph (including random regular graphs), the derivation below will be introduced using the square lattice/graph. For square lattices in dimension d , each simplex will have $2d + 1$ nodes and is denoted graphically as a star graph () with the different color combinations (e.g., , , ...) describing the various degrees of freedom for the simplex. In these simplices, the central node is the only node on which a color-switch will be attempted. For any specific set of peripheral nodes (say ) “state b” and “state r” are defined as simplices whose central node is of color “b” and “r”, respectively. So, for the peripheral config-

uration “”, “state r” is “” and “state b” is “”. There are $2d + 1$ number of simplex pairs that describe energetically unique environments, each corresponding to the a unique number n_b of “b”s that surround a central node ($n_b \in \{0, 1, \dots, 2d\}$). Note that, energetically, the exact configuration of neighbors do not matter since their energies are identical; so,  and  are equivalent simplices. Important to our treatment, each simplex pair can be treated as a single-site color switch conceptually occurring somewhere in the infinite assembly.

Establishing necessary criteria for ISD plateau boundaries. Given a sampling algorithm, inherent structures are “stuck” conformations, i.e., zero degrees of freedom are available to any site in such structures (unless temperature is added to the system). Therefore, any plateau boundary in the ISD must lie at regions of ($\epsilon_{bb}, \epsilon_{br}$)-space where at least one simplex freely switches back and forth (e.g.,  \leftrightarrow ) as such a boundary would indicate that adjacent regions in the energy space cause locally irreversible switches for that simplex ( \leftarrow  or  \rightarrow ) which is a precursor to a jammed or “stuck” conformation. Not all such moves are likely to cause jammed conformations, as peripheral sites may also be allowed to evolve to “unjam” the simplex in question. However, such equalities are necessary criteria for ISD boundaries, whose closed forms will be derived below.

For any simplex pair, one can assess whether a given energy regime would allow for free switching ( \leftrightarrow ) or switching only in one direction ( \leftarrow  or  \rightarrow ) by comparing the difference between the interaction energy of the state b simplex versus the state r simplex. In particular, the following relationship, if met, allows for free switching, and if not met results in switching permitted only in one direction:

$$\text{energy}(\text{state b}) = \text{energy}(\text{state r}) \quad (3)$$

$$\text{e.g., } \text{energy}(\text{)}) = \text{energy}(\text{)})$$

For any pair with n_b blue neighbors, and $\epsilon_{rr} \equiv 0$, this relationship becomes:

$$A\epsilon_{bb} + B\epsilon_{br} + \mu_b = C\epsilon_{bb} + D\epsilon_{br} \quad (4)$$











where μ_b is the chemical potential for a node with color ‘b’, which dictates the value of the relative fraction of ‘b’ in solution or f_b^{sol} (see Section II C); A and C are the number of ϵ_{bb} bonds in simplex states r and b, respectively, and B and D are the number of ϵ_{br} bonds in states r and b, respectively. A simplex pair with n_b number of “b” neighbors yields

$$A = n_b; \quad B = k - n_b; \quad C = 0; \quad D = n_b \quad (5)$$

See, e.g., Fig. 4a for an example of such assignments. n_b is a boundary-specific value set by the number of peripheral ‘b’s (n_b) displayed by the specific simplex pair. The additional parameter k is the degree of the regular graph (and so $k = 2d$ for any square lattice in dimension d). Applying Eqns. 5 to Eqn. 4 gives

$$\epsilon_{br} = \frac{n_b}{(2n_b - k)}\epsilon_{bb} + \frac{\mu_b}{(2n_b - k)} \quad (6)$$

a simplex pairs and their neighboring bonds

		(state b)		(state r)	
central atom color	neighbors	# of ϵ_{bb} bonds	# of ϵ_{br} bonds	# of ϵ_{bb} bonds	# of ϵ_{br} bonds
state b	state r	A	B	C	D
		0	4	0	0
		1	3	0	1
		2	2	0	2
		3	1	0	3
		4	0	0	4

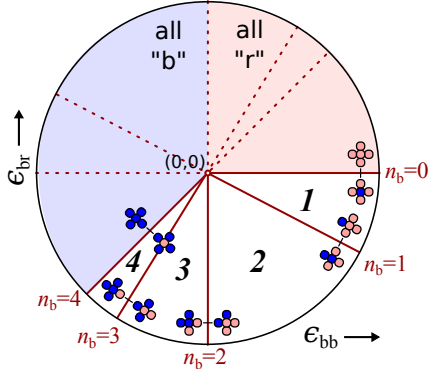
b resulting boundaries from trapped simplices

Figure 4. **Obtaining a closed-form description of plateau boundaries in $d=2$.** Plateau boundaries may be deciphered by comparing for each simplex pair (state r and b) the number of b-b and b-r bonds available to each state (a). Applying these numbers to the basic transition equation (Eqn. 4) results in borders that depend only on the number of surrounding ‘b’s (n_b), solution concentrations of each component ($\propto \mu_b$), and the connectivity ($k = 2d$) of the lattice (Eqn. 6). Plotting Eqn. 6 results in plateaux (b) that are identical to those obtained from inherent structure sampling (Fig. 2a). Correspondence in other lattices is also observed in Fig. 5.

This is the equation for all possible plateau borders within the antiferromagnetic regime of the ISD (defined by those regions in Fig. 4b whose structures are not all-red and all-blue; i.e., demarcated by simultaneous conditions $\epsilon_{br} < \epsilon_{rr} \equiv 0$ and $\epsilon_{br} < \epsilon_{bb}$). There can only be $k+1$ number of such boundaries, as n_b can be only one of $k+1$ values ($n_b = \{0, 1, \dots, k\}$). Aside from system properties (k, μ_b), each boundary is defined by the integer n_b (Fig. 4b). Additionally, a possible plateau is defined by adjacent n_b s, with the adjacent boundaries being $n_b - 1$ and n_b for $n_b \leq k$ (Fig. 4b). Attesting to the utility of the “simplex method” above, the plateau boundaries obtained from our sampling methodology are precisely marked by boundaries indicated by Eqn. 6. Compare, e.g., plateaux obtained from simulation (Fig. 2a) and Eqn. 6 in (Fig. 4b); similar comparisons for other dimensions are shown in Fig. 5.

Local structures associated with each plateau.

This section shows that the simplex formalism discussed above may be used to predict the types of inherent structures displayed by each plateau. Fig. 6 displays struc-

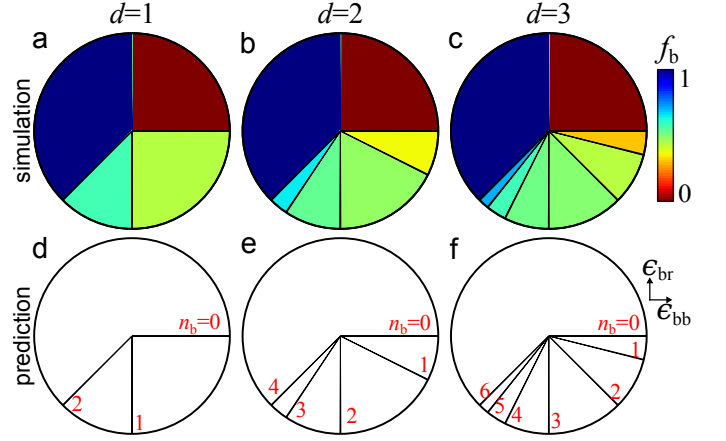


Figure 5. **Plateau boundaries obtained from simulation (a-c) and Eqn. 6 (d-f) coincide within the antiferromagnetic regime.** Simulations were performed for square lattices of varying dimension (d). For any $(\epsilon_{bb}, \epsilon_{br})$ -pair, the starting conditions were chosen to be a random mixture of equal parts blue and red and $f_b^{\text{sol}} = 0.5$ ($\mu_b = 0$); the relevant plateaux (with $0 < f_b < 1$) remain stable regardless of the starting conditions¹⁴. The antiferromagnetic regions in an ISD (a-c) display inherent structures that are neither all blue nor all red.

tural properties and a structure for each of the plateaux in $d = 2$ (Fig. 4b). For each column that describes a distinct plateau, Panel (a) describes the allowed (downhill) moves available to our sampling algorithm (at $T = 0$). Panel (b) describes the local environments displayed by the structure, as measured by f_b^i and f_r^i (these are the respective probabilities that a blue and red node is surrounded by i number of blue neighbors; $i \in \{0, \dots, k\}$). The complete collection of f_b^i and f_r^i completely describes how the structure is locally distributed. Finally, an example of the resulting polycrystalline structure is shown in Fig. 6c, and the relationships between interaction energies within each plateau is shown in Table I.

Polycrystalline types dictated by n_b . Plateaux within the antiferromagnetic regime are associated with not one structure but an ensemble of configurations that display specific distributions of local environments. Interestingly, as suggested by the plateau-specific simplex switching rules in Fig. 6a, crucial aspects of the distribution of f_b^i and f_r^i depend primarily on the characteristic n_b that describes the plateau:

$$f_b^i = 0 \text{ if } i \geq n_b \quad (7)$$

$$f_r^i = 0 \text{ if } i < n_b \quad (8)$$

These vanishing probabilities are critical rules for understanding the structural features observed in each plateau; for example, $f_b^{i \geq 1} = 0$ and $f_r^{i < 1} = 0$ for plateau 1 ($n_b = 1$) ensure the observed polycrystalline structure (Fig. 6, first column)^{14,15}, where red-blue binary crystalline domains are separated by purely red boundaries. The simplex method appears to therefore be useful to both apportion the ISD into plateaux (Figs. 4, 5), and

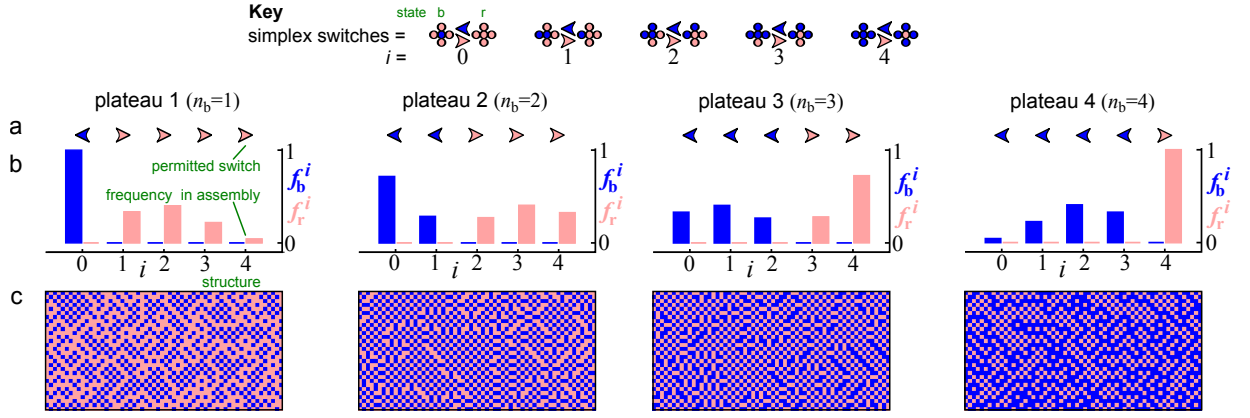


Figure 6. **Local structure evident within a plateau (b,c) are direct outcomes of energetically downhill simplex switching rules (a).** Simplex switching involves swapping the color of the central node of a simplex. The key indicates possible simplex switches, where the simplex pair is defined by a specific number of blues in the neighborhood (i ; see Definitions). Panel (a) describes the allowed moves within simplices of particular neighborhood i , where the blue and light red arrows respectively indicate a simplex switch towards state ‘b’ and ‘r’. Given that only energetically downhill switches are allowed, these unidirectional switching rules are hard constraints on the types of simplices that are allowed within an inherent structure, which is measured in (b) by f_b^i and f_r^i . (f_x^i is the number of ‘x’ sites that are surrounded by i number of blue sites divided by the total number of ‘x’ sites.) $f_r^i = 0$ if $i < n_b$ and $f_b^i = 0$ if $i \geq n_b$ where n_b is the plateau number. This, in turn, dictates the specific types of compositionally polycrystalline materials that arise from each plateau (c), which are all marked by grains composed of alternating red-blue coloring, but whose grain boundaries are plateau-specific (going from all-red for plateau 1 to all-blue plateau 4, and distinct mixtures in between). Local environments for other dimensions and lattices are shown in Fig. 7.

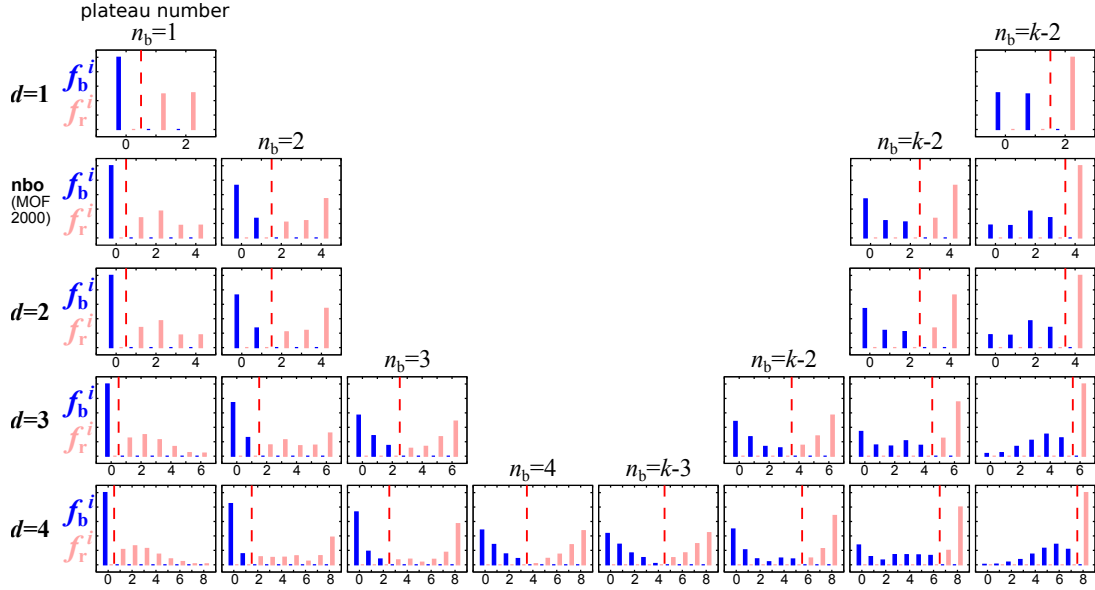


Figure 7. As shown in Fig. 6b, panels within each row describe the local environment within each ISD plateaux obtained from a particular lattice. The lattices shown are square lattices of dimension $d = \{1, 2, \dots, 5\}$ ($k = 2d$) and the **nbo** lattice ($k = 4$). These graphs display two overarching constraints that are plateau-specific: $f_b^i = 0$ if $i \geq n_b$ (see Eqn. 7) and $f_r^i = 0$ if $i < n_b$ (see Eqn. 8). These hard limits are due to the plateau number (n_b) and can be visualized within each graph (b) as a vertical barrier (red dashed line) that separates the blue bars (to the left) from the light red bars (to the right).

account for the qualitative features of the material emerging from such plateaux (Fig. 6). Given that these predictions are based on local connectivity only, our boundary predictions also hold for higher dimensions (Fig. 5).

It is important to note that all of our results are dependent on simplices, which are local snapshots of the larger

assembly. While this treatment is useful to motivate the existence of ISD plateaux and establish constraints on the local neighborhoods within plateau-specific inherent structures, it appears as though processes at the global network level dictate the *exact* magic ratios observed in each plateau. This is most evident when comparing two

Plateau #	Interaction energy constraints		
	1	2	3
1	$\epsilon_{rr} = 0$	$\epsilon_{br} < 0\epsilon_{bb}$	$-\frac{1}{2}\epsilon_{bb} < \epsilon_{br}$
2	$\epsilon_{rr} = 0$	$\epsilon_{br} < -\frac{1}{2}\epsilon_{bb}$	$0\epsilon_{br} < \epsilon_{bb}$
3	$\epsilon_{rr} = 0$	$\epsilon_{bb} < 0\epsilon_{br}$	$\epsilon_{br} < \frac{3}{2}\epsilon_{bb}$
4	$\epsilon_{rr} = 0$	$\frac{3}{2}\epsilon_{bb} < \epsilon_{br}$	$\epsilon_{br} < \epsilon_{bb}$

Table I. **Allowed interaction energy relationships within each plateau for $d=2$.** Relationships obtained from Eqn. 6 that place constraints on the inter-subunit interaction energies that account for the plateaux in Fig. 4.

networks of identical degree or coordination (k): the **nbo** lattice and $d = 2$ square lattice (rows two and three in Fig. 7), both of which have $k = 4$. While our method correctly predicts the plateau boundaries and constrains on local structure that result in the specific class of compositional polycrystal, the magic-number ratios are different in both networks (they are respectively ~ 0.3 and ~ 0.364 for the **nbo** and 2d square lattice)^{14,15}. This arises from the fact that specific distributions within the unconstrained regions of the neighborhood vary between lattice types (rows two and three in Fig. 7). This indicates that while the local connectivity describes the *type* of (polycrystalline) structure within each plateau, the actual ratio of reds:blues in solution are the result of non-local events as well; the exact effect of global connectivity on the magic number ratios is left for a future report.

Allowing solution compositions to vary. The last term containing μ_b in Eqn. 6 indicates that changing the relative implicit solution concentration of components from 1:1 (i.e., setting $\mu_b \neq 0$) will result in simply the skewing of the landscape such that the center of the radiating landscape shifts by $\mu_b/(2n_b - k)$ in the ϵ_{br} -direction and $-\mu_b/2n_b$ in the ϵ_{bb} -direction. This skew in the cen-

ter of the radial ISD versus the center of the coordinate system ($\epsilon_{bb}, \epsilon_{br} = 0, 0$) is also observed in simulations¹⁴, and indicates that, given a set of interaction energies, one may be able to predict not only the assembly’s expected plateau/inherent structure, but also the solution conditions (range of μ_b) within which an out-of-equilibrium assembly’s magic ratio will remain the same.

IV. ENDING REMARKS

Simple local rules of an interaction network, along with simple energetic considerations, have allowed for the charting of regions (plateaux) of the two-component inherent structure diagram (ISD), within which identical outcomes occur. This work allows for the characterization of the classes of inherent structures that inhabit those plateaux, and re-establishes that local connectivity (degree) explains much in the inherent structure world. However, global connectivity is finally needed to establish exactly what magic number ratios are described by the inherent structures within each plateau, which I leave open as a future line of research. This work may eventually be useful in the rational design of a range of compositional polycrystalline materials formed far from equilibrium.

ACKNOWLEDGMENTS

I thank Stephen Whitelam for discussions, and Alana Canfield Mannige and Anthony M. Frachioni for a detailed reading of the manuscript. This work was done at the Molecular Foundry at Lawrence Berkeley National Laboratory (LBNL), supported by the Office of Science, Office of Basic Energy Sciences, of the U.S. Department of Energy under Contract No. DE-AC02-05CH11231.

* rvmannige@lbl.gov

¹ Joseph D Bryngelson, Jose Nelson Onuchic, Nicholas D Socci, and Peter G Wolynes, “Funnels, pathways, and the energy landscape of protein folding: a synthesis,” *Proteins: Structure, Function, and Bioinformatics* **21**, 167–195 (1995).

² K. A. Dill and H. S. Chan, “From levinthal to pathways to funnels,” *Nat Struct Biol* **4**, 10–19 (1997).

³ J. N. Onuchic, Z. Luthey-Schulten, and P. G. Wolynes, “Theory of protein folding: the energy landscape perspective,” *Annu Rev Phys Chem* **48**, 545–600 (1997).

⁴ Garegin A. Papoian and Peter G. Wolynes, “The physics and bioinformatics of binding and folding—an energy landscape perspective,” *Biopolymers* **68**, 333–349 (2003).

⁵ Jin Wang and Gennady M Verkhivker, “Energy landscape theory, funnels, specificity, and optimal criterion of biomolecular binding,” *Physical review letters* **90**, 188101 (2003).

⁶ Hung D Nguyen, Vijay S Reddy, and Charles L Brooks III, “Invariant polymorphism in virus capsid assembly,” *Journal of the American Chemical Society* **131**, 2606–2614 (2009).

⁷ José D Faraldo-Gómez and Benoît Roux, “On the importance of a funneled energy landscape for the assembly and regulation of multidomain src tyrosine kinases,” *Proceedings of the National Academy of Sciences* **104**, 13643–13648 (2007).

⁸ Peter E Leopold, Mauricio Montal, and José N Onuchic, “Protein folding funnels: a kinetic approach to the sequence-structure relationship,” *Proceedings of the National Academy of Sciences* **89**, 8721–8725 (1992).

⁹ Mark T Oakley, David J Wales, and Roy L Johnston, “Energy landscape and global optimization for a frustrated model protein,” *J Phys Chem B* **115**, 11525–11529 (2011).

¹⁰ Cade B Markegard, Iris W Fu, K Anki Reddy, and Hung D Nguyen, “Coarse-grained simulation study of sequence effects on dna hybridization in a concentrated environment,”

- The Journal of Physical Chemistry B **119**, 1823–1834 (2015).
- ¹¹ Lester O Hedges, Ranjan V Mannige, and Stephen Whitelam, “Growth of equilibrium structures built from a large number of distinct component types,” *Soft matter* **10**, 6404–6416 (2014).
 - ¹² Anthony J Kim, Raynaldo Scarlett, Paul L Biancaniello, Talid Sinno, and John C Crocker, “Probing interfacial equilibration in microsphere crystals formed by dna-directed assembly,” *Nature materials* **8**, 52–55 (2009).
 - ¹³ Raynaldo T Scarlett, John C Crocker, and Talid Sinno, “Computational analysis of binary segregation during colloidal crystallization with dna-mediated interactions,” *The Journal of chemical physics* **132**, 234705 (2010).
 - ¹⁴ Andrew C.-H. Sue, Ranjan V. Mannige, Hexiang Deng, Dennis Cao, Cheng Wang, Felipe Gándaraa, Fraser Stoddart, Stephen Whitelamc, and Omar M. Yaghi, “Two-component metal-organic framework displaying compositional robustness to solution constitution,” *Proceedings of the National Academy of Sciences* **112**, 5591–5596 (2015).
 - ¹⁵ R. V. Mannige and S. Whitelam, “Predicting the outcome of the growth of binary solids far from equilibrium,” Submitted to the PRE; arXiv:1507.06267 [cond-mat.soft] (2015).
 - ¹⁶ Eduardo Sanz, Chantal Valeriani, Daan Frenkel, and Marjolein Dijkstra, “Evidence for out-of-equilibrium crystal nucleation in suspensions of oppositely charged colloids,” *Physical review letters* **99**, 055501 (2007).
 - ¹⁷ Baron Peters, “Competing nucleation pathways in a mixture of oppositely charged colloids: Out-of-equilibrium nucleation revisited,” *The Journal of chemical physics* **131**, 244103 (2009).
 - ¹⁸ Stephen Whitelam, Rebecca Schulman, and Lester Hedges, “Self-assembly of multicomponent structures in and out of equilibrium,” *Physical review letters* **109**, 265506 (2012).
 - ¹⁹ Stephen Whitelam, Lester O Hedges, and Jeremy D Schmit, “Self-assembly at a nonequilibrium critical point,” *Physical review letters* **112**, 155504 (2014).
 - ²⁰ Stephen Whitelam and Robert L Jack, “The statistical mechanics of dynamic pathways to self-assembly,” *Annu Rev Phys Chem.* **66**, 143–163 (2015).
 - ²¹ Frank H Stillinger and Thomas A Weber, “Inherent structure in water,” *The Journal of Physical Chemistry* **87**, 2833–2840 (1983).
 - ²² FH Stillinger, “Dynamical lattice gases and inherent structure in condensed phases,” *The Journal of chemical physics* **88**, 380–385 (1988).
 - ²³ Eric Bertin, “Real-space analysis of inherent structures,” *EPL (Europhysics Letters)* **71**, 452 (2005).
 - ²⁴ N Norman, N Greenwood, and A Earnshaw, “Chemistry of the elements,” Butter worth-Heinemann, Earnshaw, Alan, (1997).
 - ²⁵ Wikipedia URL:
https://en.wikipedia.org/wiki/Niobium_monoxide.